

REMARKS

The present application was filed on December 14, 1999 with claims 1-8. In the outstanding Office Action dated December 7, 2000, the Examiner has: (i) rejected claims 1 and 3-5 under 35 U.S.C. §102(e) as being anticipated by U.S. Patent No. 6,097,806 to Baker et al. (hereinafter “Baker”); (ii) rejected claim 6 under 35 U.S.C. §103(a) as being unpatentable over Baker; and (iii) rejected claims 2, 7 and 8 under 35 U.S.C. §103(a) as being unpatentable over Baker in view of U.S. Patent No. 5,845,282 to Alley et al. (hereinafter “Alley”).

In this response, Applicants have amended claim 1 for clarity, presented additional claims 9-19, and traverse the §102(e) and §103(a) rejections for at least the reasons set forth below. It is asserted that no new matter has been introduced by the amendments made herein. Applicants respectfully request reconsideration of the present application in view of the following remarks.

The present invention relates to “methods and apparatus for contingent transfer and execution of spoken language interfaces” (Specification; page 2, lines 24-25). In a portable speech assistant (PSA), “a spoken language interface is defined in sets of user interface files. These are referred to as vocabularies files, prompt files, profiles and scripts depending on the role they play in structuring the interface” (Specification; page 3, lines 3-5). As used by the present invention, the term “spoken language interface” is intended to refer to the general act of speaking to a machine, listening to a machine, and/or interacting with a machine through utterances or audible expressions, and does **not** refer to a particular lingual type (e.g., English or Spanish).

An important aspect of the present invention is its ability to dynamically instantiate a new application and its spoken language interface (Specification; page 42, lines 5-7). It should be appreciated that the spoken language user interface is a collection of operable features that allows a user to interact with the application. For example, user utterances may operate the features of the application, e.g., by supplying a reference to one or more events to be processed by the target application (Specification; page 3, lines 5-6). The term “event” is used by the present invention in a conventional sense in the context of event handling programs. Event handling is a feature of the

application. These operable features, which are built into an application and are controlled at least in part by user utterances, are to be distinguished from data on which the application program acts.

Claims 1 and 3-5 stand rejected under 35 U.S.C. §102(e) as being anticipated by the Baker reference. With specific regard to independent claim 1, the Examiner contends that “Baker teaches, ‘In an apparatus for a portable language interface for a user to a device in communication with the apparatus, . . . a method for modifying a data structure containing the at least one interface data,’ comprising: ‘adding a new application to the device’ (col. 5, lines 33-55); ‘generating a second user interface data set in accordance with the new application’ (col. 5, lines 33-55); ‘transferring the second user interface data set from the device to the apparatus and loading the second user interface data set into the data structure of the apparatus’ (col. 6, lines 22-34).” Applicants respectfully disagree with this contention.

The Baker reference is directed to a method and apparatus “for selecting a language for use by a plurality of [human] agents” in the context of an automatic call distribution (ACD) system (Baker; column 2, lines 49-51), and more specifically relates to presenting text to human agents of the automatic call distributor (Baker; column 1, lines 6-8). Baker discloses that “[b]ecause of the many multi-lingual uses for ACDs a need exists for a way of easily adapting ACDs for use with other languages” (Baker; column 2, lines 44-46). In Baker, it is assumed that an enterprise exists which receives telephone calls from one or more customers, each of which may speak a different language, and that the terminal of the agent servicing a call should present the agent with his or her script of things to say to a customer, in the language of the particular customer, in essence providing a translation (Baker; column 2, lines 14-57). In contrast to the present application, Baker uses the term “language” to denote a particular lingual type, such as English or Spanish. For example, Baker states that “[t]he ACD may be setup for use with any particular language (*e.g., German*) by a supervisor (not shown) entering a default language indicator through a supervisor’s terminal” (Baker; column 5, lines 21-24; emphasis added).

Applicants respectfully assert that claim 1 is patentable over the Baker reference. Specifically, Baker fails to teach or remotely suggest, among other things, interaction with an

application program at least in part via user utterances in conjunction with a spoken language interface. As noted above, Baker uses the term “language” to refer to a certain lingual type (e.g., English or Spanish) rather than to the act of speaking. Baker provides messages to a user (agent) in a particular national language so that the agent can then speak with a customer in the selected language. As Baker states, “[t]ext for sales presentations to customers may be presented to the agent which may be read to the calling party... [T]he ACD 10 may present messages to agents under any of a number of languages... (e.g., German)” (Baker; column 5, lines 17-25). It is important to understand, however, that Baker does not interact with the ACD application by way of a spoken language interface. In fact, essentially the only user interaction taught by Baker is choosing the type of language in which to display text to the agents.

Additionally, Baker fails to teach or suggest “adding a new application to the device” as defined by claim 1 of the present invention. Other than the ACD application itself, “programming the file through the supervisor’s terminal 14 with the language indicator” (Baker; column 5, lines 35-37) is not to be equated with adding a new application. Rather, Baker merely discloses modifying a data item used by the agent terminal. An application as used in the conventional sense, and as intended by the present application, refers to executable software. Consequently, downloading “the attribute file and language indicator ... to the individual agent stations” as taught by Baker (Baker; column 5, lines 39-41) is to be clearly distinguished from “adding a new application” as defined in claim 1 of the present application. Similarly, “messages transferred from the CPU 46 to the BRL” (Baker; column 5, lines 50-51) are not considered applications.

Baker also fails to teach or suggest “generating a second user interface data set in accordance with the new application” as defined by claim 1. The Examiner contends that this step is disclosed in Baker at column 5, lines 33-55. In disagreement, however, Applicants respectfully assert that “messages (displayed to the agents) into a language consistent with the language indicator” (Baker; column 5, lines 49-50) are merely textual instructions for responding to customers, expressed in a national language, and not a user interface to the computing apparatus as the term is defined by the present invention. The user interface data set, as recited in claim 1 of the present application,

comprises “spoken language interface elements and data recognizable by the application of the device.” Notwithstanding the above traversal, Applicants have amended claim 1 to more clearly define the second user interface data set as “representing spoken language interface elements and data recognizable by the new application.” This language is similarly stated in the preamble of claim 1. These features are not taught or suggested by the Baker reference.

For at least the foregoing reasons, Applicants respectfully submit that claim 1 is patentable over the Baker reference. Accordingly, favorable reconsideration and allowance of claim 1 is respectfully solicited.

With respect to claims 2-5, which depend from claim 1, Applicants respectfully assert that these claims are also patentable over the prior art of record by virtue of their dependency from claim 1, which is believed to be patentable for at least the reasons set forth above. Moreover, these claims define additional patentable subject matter in their own right. For example, claim 3 further defines the invention as including the step of “removing a user interface data set from the data structure.” Since the Baker reference fails to teach or suggest any user interface data set, Applicants submit that these claims are patentable over the cited prior art, not merely by virtue of their dependency from claim 1, but also in their own right. Accordingly, favorable reconsideration and allowance of claims 2-5 is respectfully requested.

Claim 6 stands rejected under 35 U.S.C. §103(a) as being unpatentable over the Baker reference. In paragraph 4 of the present Office Action, the Examiner acknowledges that “Baker does not explicitly teach spoken language” but “[i]t would have been obvious to one of ordinary skill in the art at the time of the invention to use spoken language data instead of language data so as to create correct pronunciation of that language.” In this regard, Applicants respectfully disagree with the Examiner’s characterization of the present invention.

Specifically, Baker addresses a completely different problem than that of the present invention, namely, providing textual messages to agents in a selected national language in the context of an automatic call distribution (ACD) system (Baker; Abstract). As noted above, Baker teaches that “[b]ecause of the many multi-lingual uses for ACDs a need exists for a way of easily

adapting ACDs for use with other languages” (Baker; column 2, lines 44-46). Baker’s solution, addressing the human-human interface, fails to address the problem defined by the present invention, nor can it be modified to solve that problem, namely, providing “spoken language interface methods and apparatus” for updating the man-machine interface based on changes in an application or engine (Specification; page 2, lines 17-21; emphasis added).

As stated above in connection with claim 1, the use of the term “language” in Baker is distinguishable from the manner in which that term is defined by the present application. As used by the present invention, the term “spoken language interface” is intended to refer to the general act of communicating between human and machine through spoken utterances or audible expressions, and does not refer to a particular lingual type (e.g., English or Spanish), as in Baker. In this regard, Baker does not teach or suggest a spoken language interface. Furthermore, the present invention does not use spoken language data “to create correct pronunciation of that language,” as the Examiner contends. Instead, such spoken language data is employed by the present invention to interface with an application and address one or more operable features of the application in response thereto. Consequently, since the Baker reference does not teach or remotely suggest these features as claimed by the present application, Applicants submit that claim 6 is patentable over the prior art of record and therefore favorable reconsideration and allowance of this claim is respectfully requested.

With regard to claims 7 and 8, which depend from claim 6, Applicants respectfully assert that these claims are also patentable over the prior art of record by virtue of their dependency from claim 6, which is believed to be patentable for at least the reasons set forth above. Furthermore, these claims define additional patentable subject matter in their own right. Accordingly, favorable reconsideration and allowance of claims 7 and 8 is respectfully solicited.

Applicants respectfully submit that newly presented claims 9-12, which depend from claim 1, and claims 13-15, which depend from claim 6, are also patentable over the prior art of record by virtue of their dependency from their respective claims, which are believed to be patentable for at least the reasons set forth above. Moreover, these claims define additional patentable subject matter

in their own right. For example, claim 9 further defines the new application as comprising “a speech aware application, the speech aware application being responsive to user utterances for at least partially interacting with the new application.” Support for this claim is found in the present specification, at least on page 39, lines 1-8. Similarly, claims 10-15 further define the invention as comprising a prompting feature which is described at least on page 29, lines 19-25 and page 35, line 26 to page 36, line 11 of the present specification. This prompting feature is not taught or suggested by the prior art cited by the Examiner or known to Applicants. Inasmuch as the cited prior art references, when considered individually or in combination, fail to teach or suggest these additional features of the present invention, Applicants assert that claims 9-15 are patentable over the prior art, not merely by virtue of their dependency from their respective claims, but also in their own right. Accordingly, favorable consideration and allowance of these claims is respectfully requested.

Independent claims 16 and 19 have been added which are of similar scope to claim 6, claim 16 being directed to an apparatus and claim 19 being directed to an article of manufacture. For at least the reasons set forth above in connection with claim 6, Applicants respectfully assert that these claims are also patentable over the prior art of record .

With regard to newly presented claims 17 and 18, which depend from claim 16, Applicants respectfully assert that these claims are patentable over the prior art of record by virtue of their dependency from claim 16, which again is believed to be patentable for at least the reasons set forth above with respect to claim 6. Moreover, these claims define additional patentable subject matter in their own right. Accordingly, favorable consideration and allowance of claims 16-19 is respectfully requested.

Application Serial No. 09/460,913

In view of the foregoing, Applicants believe that pending claims 1-19 are in condition for allowance and respectfully request withdrawal of the § 102(e) and § 103(a) rejections.

Attached hereto is a marked-up version of the changes made to the claims by the present Amendment. The attached page is captioned "Version with markings to show changes made."

Respectfully submitted,

A handwritten signature in black ink, reading "Wayne L. Ellenbogen", with a long horizontal flourish extending to the right.

Date: March 7, 2001

Wayne L. Ellenbogen  
Attorney for Applicant(s)  
Reg. No. 43,602  
Ryan, Mason & Lewis, LLP  
90 Forest Avenue  
Locust Valley, NY 11560  
(516) 759-7662

VERSION WITH MARKINGS TO SHOW CHANGES MADE

IN THE CLAIMS:

Claim 1 has been amended as follows:

1. (Amended) In apparatus for providing a portable spoken language interface for a user to a device in communication with the apparatus, the device having at least one application associated therewith, the spoken language interface apparatus comprising: (A) an audio input system for receiving speech data provided by the user; (B) an audio output system for outputting speech data to the user; (C) a speech decoding engine for generating a decoded output in response to spoken utterances; (D) a speech synthesizing engine for generating a synthesized speech output in response to text data; (E) a dialog manager operatively coupled to the device, the audio input system, the audio output system, the speech decoding engine and the speech synthesizing engine; and (F) at least one user interface data set operatively coupled to the dialog manager, the user interface data set representing spoken language interface elements and data recognizable by the application of the device; wherein: (i) the dialog manager enables connection between the input audio system and the speech decoding engine such that the spoken utterance provided by the user is provided from the input audio system to the speech decoding engine; (ii) the speech decoding engine decodes the spoken utterance to generate a decoded output which is returned to the dialog manager; (iii) the dialog manager uses the decoded output to search the user interface data set for a corresponding spoken language interface element and data which is returned to the dialog manager when found; (iv) the dialog manager provides the spoken language interface element associated data to the application of the device for processing in accordance therewith; (v) the application of the device, on processing that element, provides a reference to an interface element to be spoken; (vi) the dialog manager enables connection between the audio output system and the speech synthesizing engine such that the speech synthesizing engine which, accepting data from that element, generates a



synthesized output that expresses that element; and (vii) the audio output system audibly presenting the synthesized output to the user; a method for modifying a data structure containing the at least one user interface data set, comprising:

adding a new application to the device;

generating a second user interface data set in accordance with the new application, the second user interface data set representing spoken language interface elements and data recognizable by the new application;

transferring the second user interface data set from the device to the apparatus; and

loading the second user interface data set into the data structure of the apparatus.

Claims 9-19 have been added as follows:

--9. (New) The method of claim 1, wherein the new application comprises a speech aware application, the speech aware application being responsive to user utterances for at least partially interacting with the new application.

10. (New) The method of claim 1, further comprising the step of:

the device prompting the user for information comprising a spoken utterance, the device manager being responsive to the spoken utterance for operatively modifying at least one of a predetermined parameter of the device and an application running on the device.

11. (New) The method of claim 10, wherein the step of prompting the user for information includes the steps of:

storing one or more user experience parameters corresponding to a familiarity of the user with a predetermined procedure of the application; and

selecting a prompt from a set of prompts for presentation to the user, the set of prompts including varying amounts of instruction based at least in part on experience parameters,

the selected prompt substantially matching the stored experience parameters of the user.

12. (New) The method of claim 10, wherein the step of prompting the user for information includes the steps of:

storing an internal data set including at least one of a date, a time and a number of times which a predetermined procedure of an application is performed; and

selecting a prompt from a set of prompts for presentation to the user, the set of prompts including varying amounts of instruction based at least in part on information included in the internal data set, the selected prompt substantially matching the stored internal data set.

13. (New) The method of claim 6, further comprising the step of:

the portable spoken language interface device prompting the user for information comprising a spoken utterance, the device being responsive to the spoken utterance for operatively modifying at least one of a predetermined parameter of the device and an application running on the device.

14. (New) The method of claim 13, wherein the step of prompting the user for information includes the steps of:

storing one or more user experience parameters corresponding to a familiarity of the user with a predetermined procedure of the application; and

selecting a prompt from a set of prompts for presentation to the user, the set of prompts including varying amounts of instruction based at least in part on experience parameters, the selected prompt substantially matching the stored experience parameters of the user.

15. (New) The method of claim 13, wherein the step of prompting the user for information includes the steps of:

storing an internal data set including at least one of a date, a time and a number of times which a predetermined procedure of an application is performed; and

selecting a prompt from a set of prompts for presentation to the user, the set of prompts including varying amounts of instruction based at least in part on information included in the internal data set, the selected prompt substantially matching the stored internal data set.

16. (New) Apparatus for automatically providing contingent transfer and execution of one or more spoken language interfaces for a user with respect to at least one external network with which the user interacts, the apparatus comprising:

a portable spoken language interface device; and

a personal data assistant (PDA) operatively coupled to the spoken language interface device, the PDA including at least one application associated therewith;

wherein the portable spoken language interface device is operative to: (i) request a spoken language interface data set from the external network upon discovery of the external network; (ii) receive from the external network the spoken language interface data set; and (iii) load the spoken language interface data set into the data structure of the portable spoken language interface device for use by the user interfacing with the external network.

17. (New) The apparatus of claim 16, wherein the portable spoken language interface device is in wireless communication with the external network.

18. (New) The apparatus of claim 16, wherein the portable spoken language interface device comprises a personal speech assistant (PSA), the PSA comprising:

an audio input system for receiving speech data provided by the user;

an audio output system for outputting speech data to the user;

a speech decoding engine for generating a decoded output in response to spoken

utterances;

a speech synthesizing engine for generating a synthesized speech output in response to text data;

a dialog manager operatively coupled to the device, the audio input system, the audio output system, the speech decoding engine and the speech synthesizing engine; and

at least one user interface data set operatively coupled to the dialog manager, the user interface data set representing spoken language interface elements and data recognizable by the application of the device;

wherein:

the dialog manager enables connection between the input audio system and the speech decoding engine such that the spoken utterance provided by the user is provided from the input audio system to the speech decoding engine;

the speech decoding engine decodes the spoken utterance to generate a decoded output which is returned to the dialog manager;

the dialog manager uses the decoded output to search the user interface data set for a corresponding spoken language interface element and data which is returned to the dialog manager when found;

the dialog manager provides the spoken language interface element associated data to the application of the device for processing in accordance therewith;

the application of the device, on processing that element, provides a reference to an interface element to be spoken;

the dialog manager enables connection between the audio output system and the speech synthesizing engine such that the speech synthesizing engine which, accepting data from that element, generates a synthesized output that expresses that element; and

the audio output system audibly presents the synthesized output to the user.

19. (New) An article of manufacture for automatically providing a spoken language interface for a user with respect to at least one external network with which the user interacts, wherein the user possesses a portable spoken language interface device having a data structure for storing one or more user interface data sets used to provide one or more spoken language interfaces, comprising a machine readable medium containing one or more programs which when executed implement the steps of:

requesting a spoken language interface data set from the external network upon discovery of the network;

transferring the spoken language interface data set from the external network to the device; and

loading the spoken language interface data set into the data structure of the device for use by the user interfacing with the external network.--